Presidential Advisory Committee Briefing

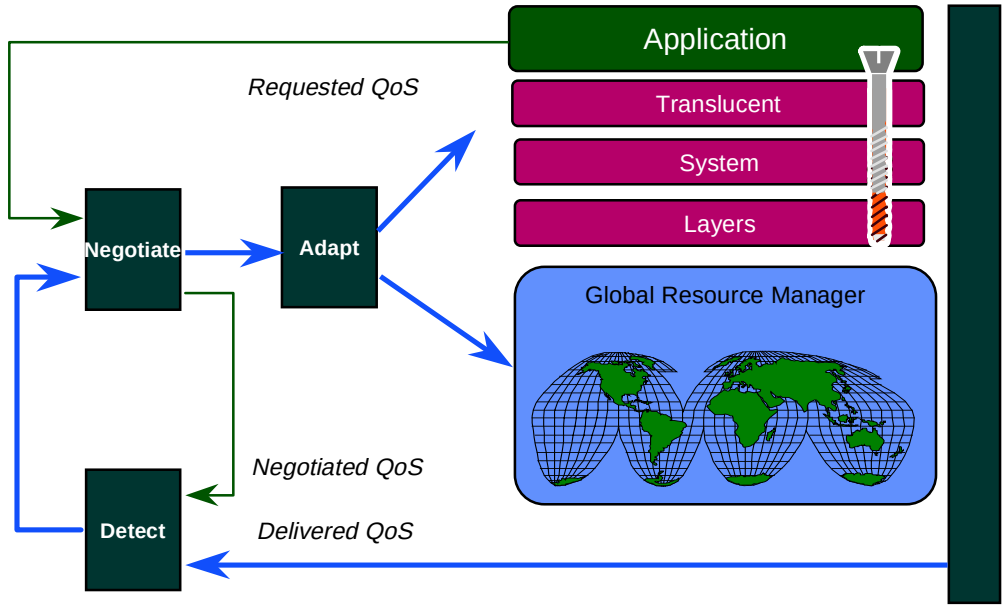
December 9, 1997

Dr. Robert F. Lucas
Defense Advanced Research Projects Agency
Information Technology Office

- Middleware and Other Software
  - Baseware
  - Protocols and Architectures
  - Software Development and Management Tools
  - Data and Information Management
  - Interoperability
  - Performance Monitoring
  - Human Interfaces
- Devices and Information/Computing Architectures
  - New Kinds of Devices and Sensors
  - End-to-End System Architecture (I.e., Clusters)
  - Intelligent Device Proliferation
  - System Support for Roaming

Quorum

Develop a global operating system that allows Defense applications to achieve predictable, controllable end-to-end quality of service on a globally managed pool of geographically distributed resources



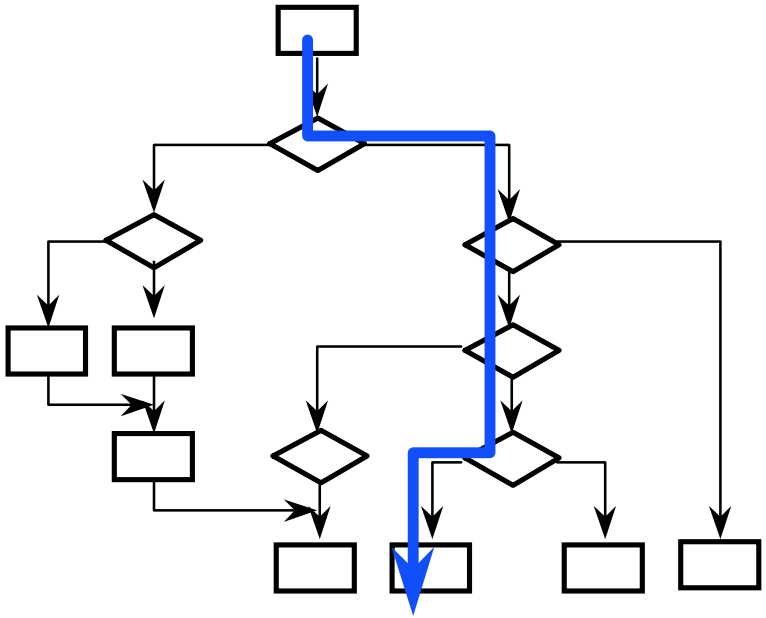
Environment

- Joint & coalition missions demand dynamic sharing of assets and adaptation to changing conditions
- Defense missions must rely on COTS technologies & an unpredictable infrastructure
- Current approaches to distributed computing do not support predictable quality of service or assurance of mission-critical properties
- No single COTS vendor can provide end-to-end solution

Specialization Toolkit (OGI)

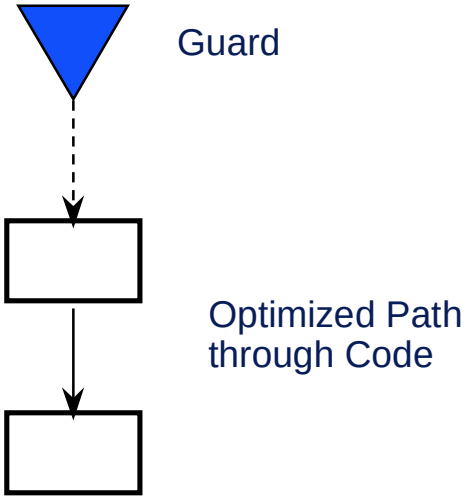
Systems Code is branch intensive

- Conditions, parameters must be checked on every invocation
- If conditions vary slowly, path taken will often remain invariant
- High overhead incurred



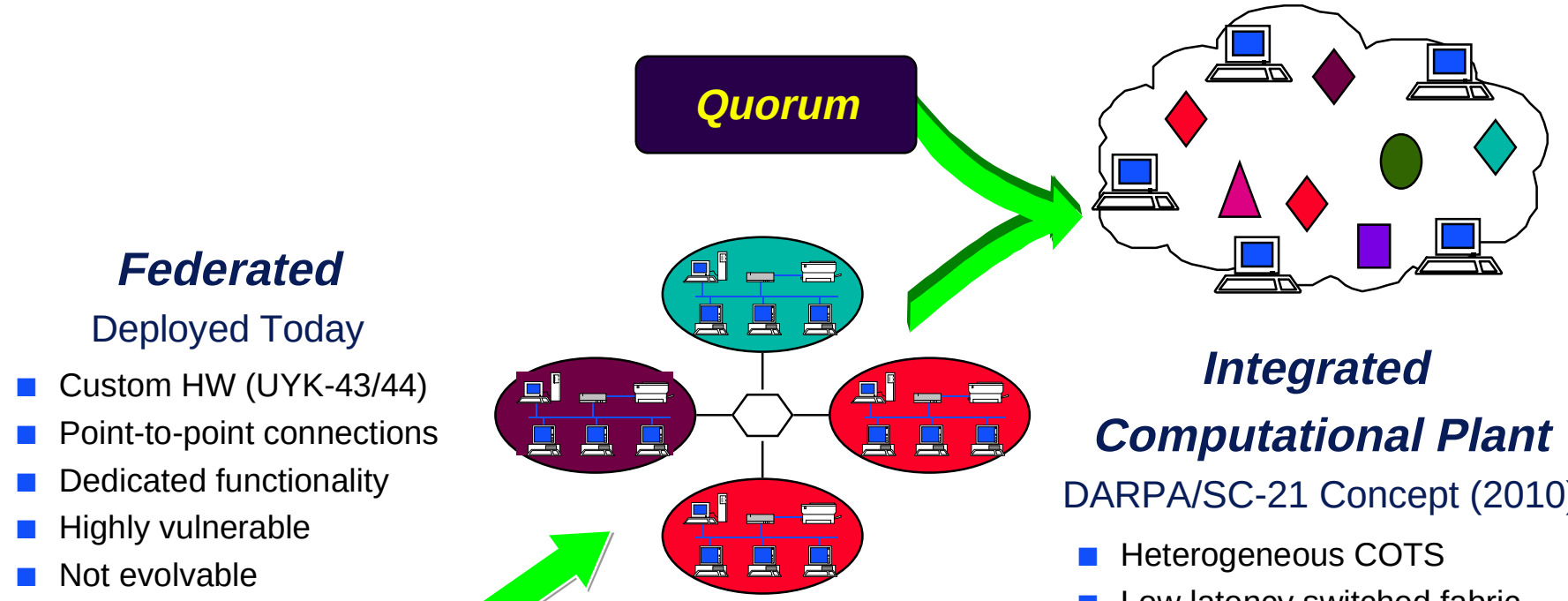
Specialization optimizes path

- Eliminate branches, evaluate constants
- Guard specialized code with simple predicate that indicates domain of validity



Quorum Insertion Opportunity:

SC-21/AdCon-21 Program (NSWC-Dahlgren)



Federated

Deployed Today

- Custom HW (UYK-43/44)
- Point-to-point connections
- Dedicated functionality
- Highly vulnerable
- Not evolvable

Distributed LAN

Aegis Baseline 7 (1998)

- Homogeneous COTS
- Network of LANs
- Fixed allocation
- Vulnerable to local damage
- Limited evolvability

Integrated Computational Plant

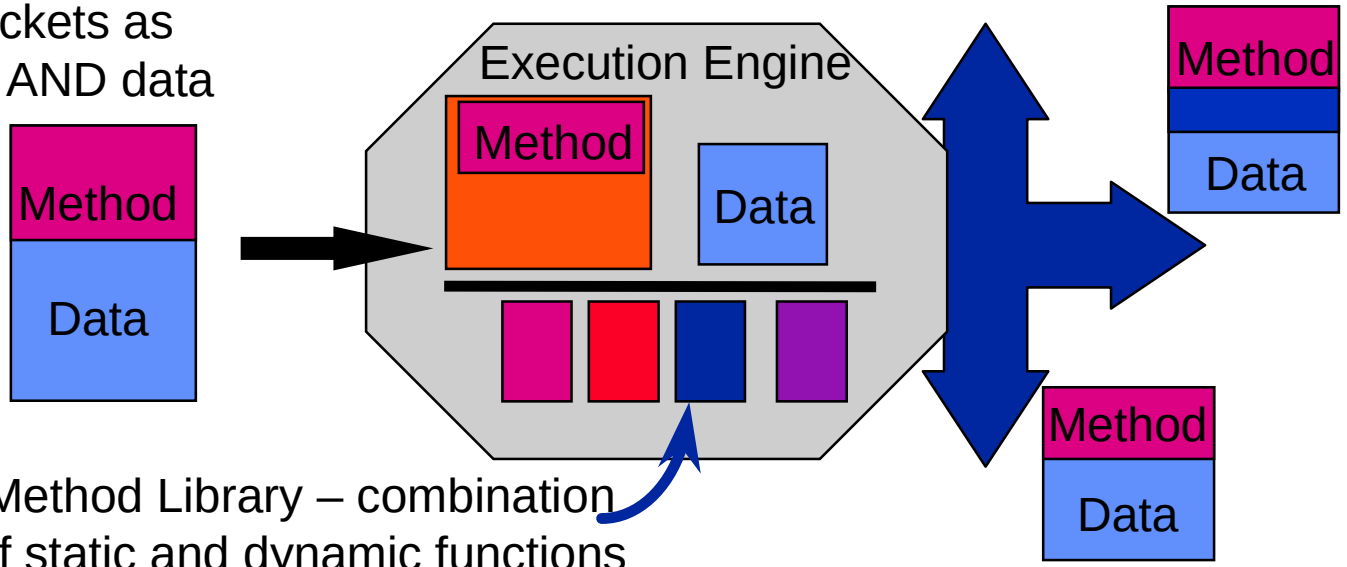
DARPA/SC-21 Concept (2010)

- Heterogeneous COTS
- Low latency switched fabric
- Dynamic allocation
- Integrated functionality
- Effectively invulnerable to localized damage
- Evolvable

Smart Packets

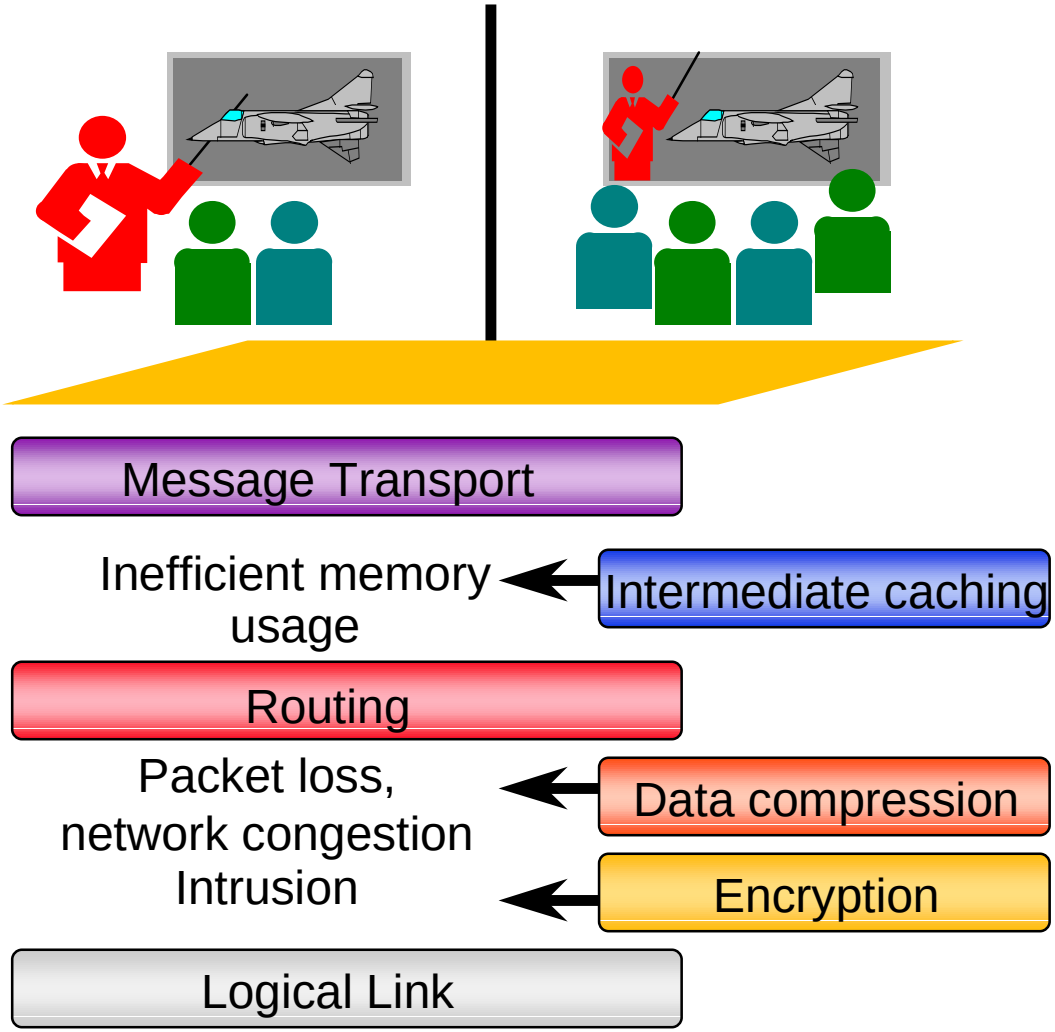
Static Packets: Network elements constrained to simple functions

Active nodes use SmartPackets as software AND data



Method Library – combination of static and dynamic functions

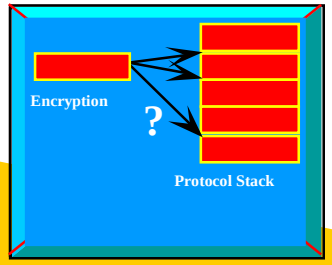
Teleconferencing Improvements During Live Sessions



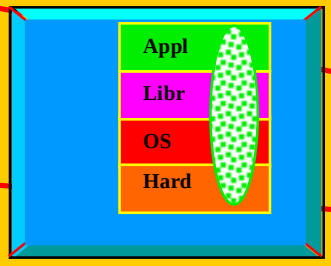
- **Active networks can counter anomalies during live sessions by introducing improvements directly into the underlying network elements**
- **The enhancements affect the intermediate processing nodes and target the physical elements closest to the problem**
- **Users see immediate qualitative improvements in teleconferencing session – clearer audio, smoother video**
- **Stronger security assurances**

Advantages of Composable Service Architecture

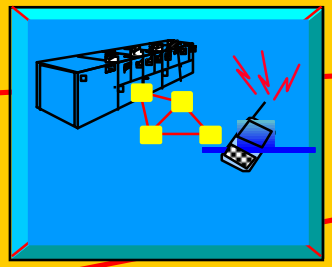
New functions can be added anywhere through active packets and composed service architecture



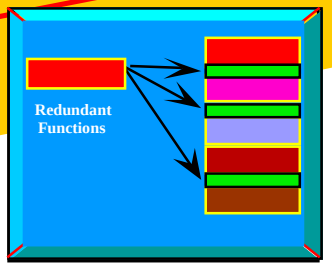
Network code at any level can be tailored exactly to application requirements



New technology integration accomplished through active injection of new services



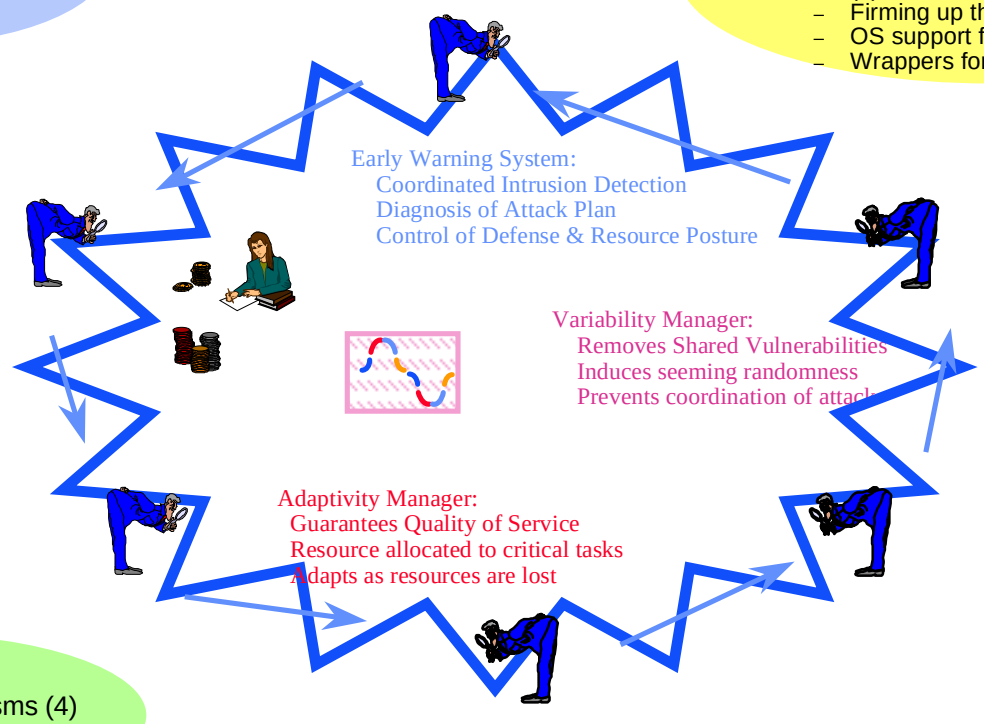
Function duplication can be avoided by runtime analysis of service stack



Information Survivability: Priorities

- **Early Warning System**
  - Intrusion Detection (1)
  - Intrusion Reponse (3)

- **Barriers**
  - Policy driven Access Control (1)
  - Operating System support for Security (1)
  - Firming up the Network Infrastructure (1)
  - OS support for Fault Tolerance (4)
  - Wrappers for Legacy Components (2)



- **Variability (5)**

- **Adaptivity**
  - Market mechanisms (4)
  - Object services to support adaptivity (4)

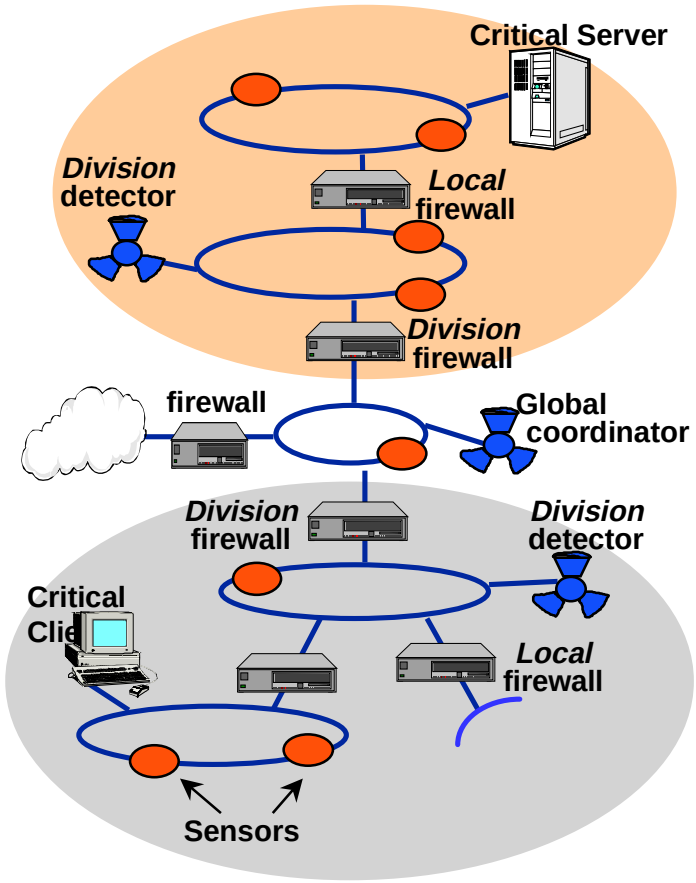
Adaptivity and variability are seminal activities. This program will allow us to understand the opportunities. Breakthrough results will occur in a follow-on program.

Representative Projects

Survivability for Large Scale Systems

Detect, isolate, reconfigure, repair

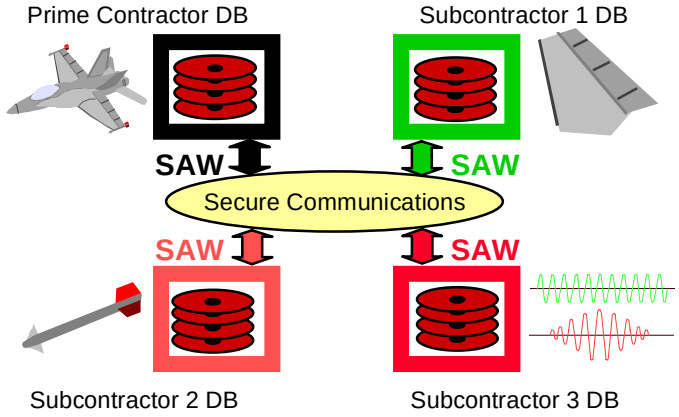
Dynamic Response (Boeing)



Wrappers and Composition

Easy insertion of barriers to attack into legacy systems

Secure Access Wrapper (SRI)



Query-Based Access Control:

Securing Oracle databases via access mediation

- Query and answer filtering
- Internal access unaffected
- External access secured
- Automated resolution of security policy conflicts

Impact (for F-22 IWSDB Case Study):

–Elimination of paper-based methods for approval of external info requests

–Potential for order-of-magnitude reduction of planning time in large-scale military operations

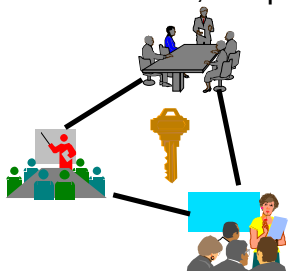
Representative Projects

High Confidence Networking

Strong barriers to network attacks

Group Communication and Key Agreement (TIS)

Challenge: Key revocation and dynamic rekeying in milliseconds, scalable to 100,000 participants



Group key changes when membership changes

1. Each organization specifies policy

- Authorizations (identity-, group-, or role-based)
- Precedence among authorizations
- Rules to deconflict rules

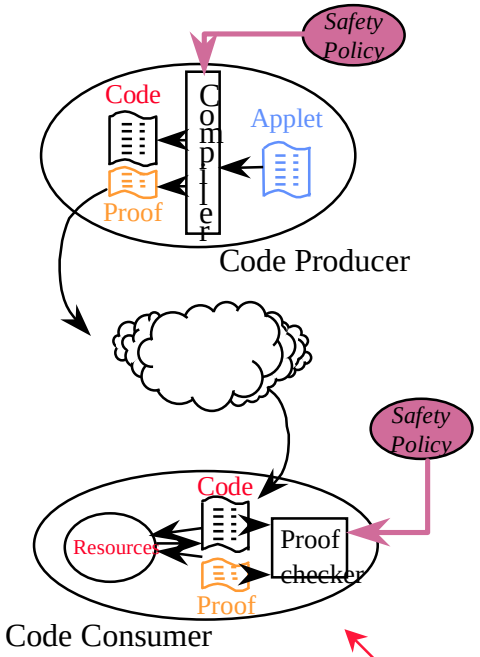
2. Negotiate a group "security context"

- security services (authentication, non-repudiation, integrity, confidentiality)
- communication layer for the service
- encryption algorithms
- key exchange algorithms
- security parameters (key lengths, key lifetimes)

High Confidence Computing

Strong barriers to attack integral to computing systems

Proof-Carrying Code (CMU)



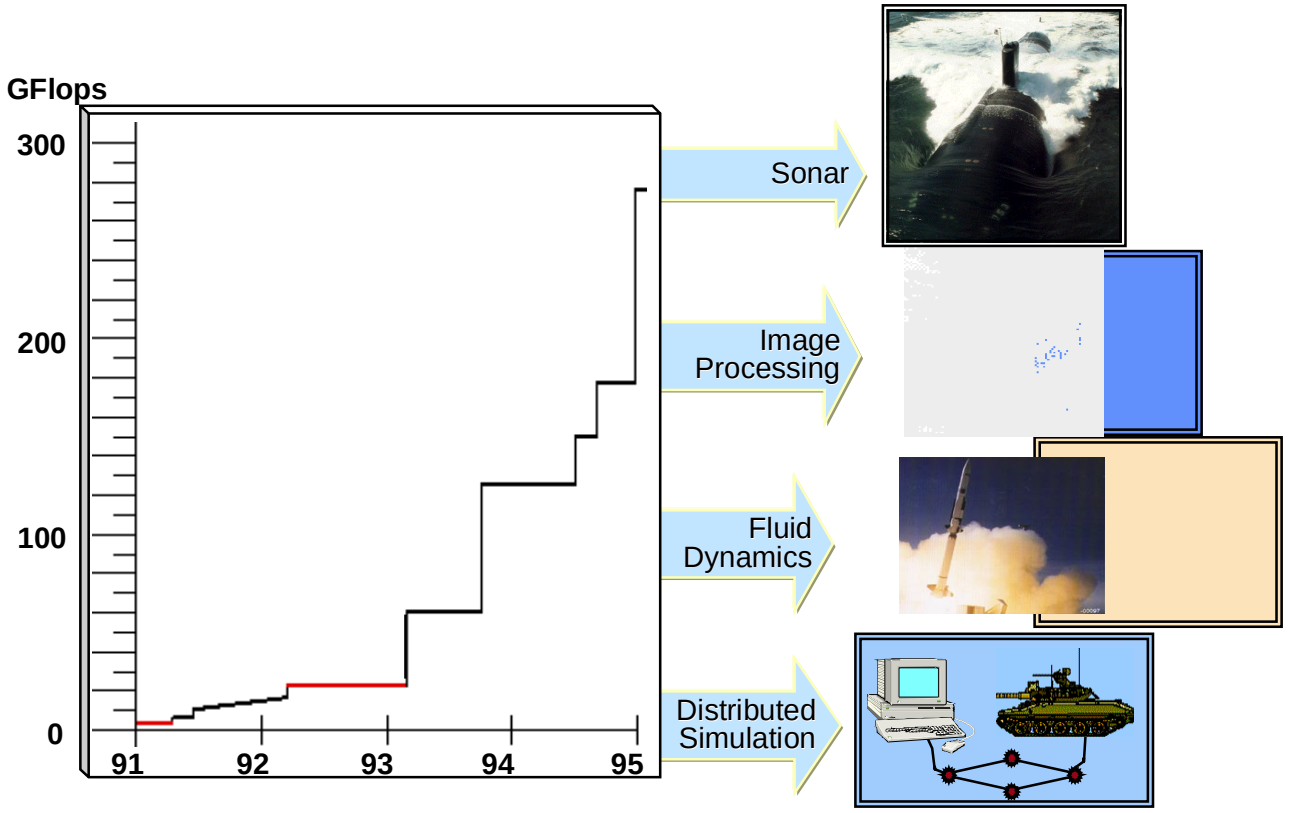
The only trusted component

Systems Environments

This program develops software technology which enables computationally intensive defense applications to effectively exploit multiprocessing systems

- ***Multiprocessing scalable hardware technology offers tremendous price/performance advantages, but...***

  ... Utility limited by lack of applications software



Environment

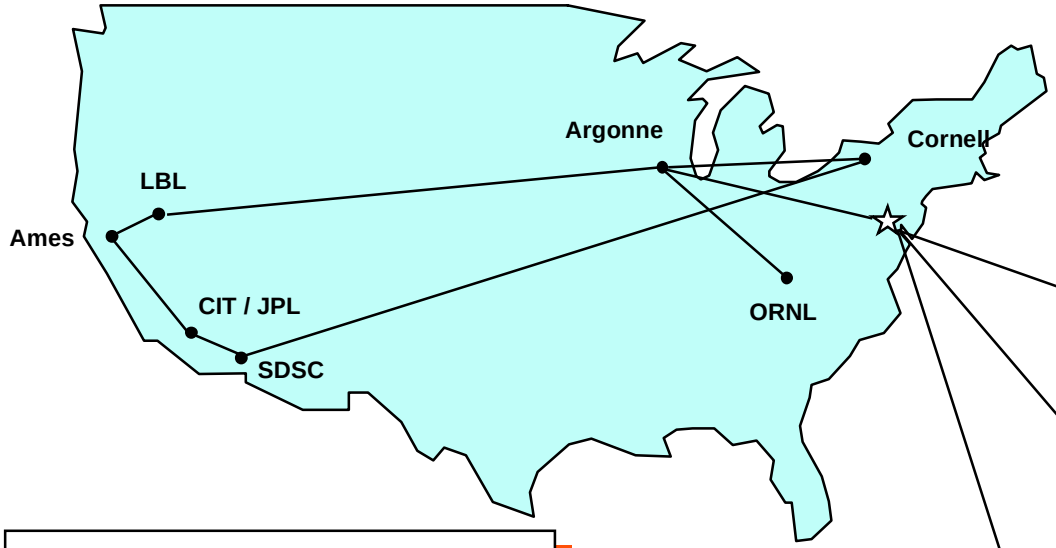
- Compute and data intensive defense applications
- Scalable multiprocessing systems
- Large existing investment in software that was not written to exploit scalable systems
- Long software system lifetimes
- Expensive to rewrite applications from scratch
- Commercial software base not as mature

Research Efforts

- Languages and Runtime support
- Scalable Libraries
- Experimental Applications

SF EXPRESS

Scalable Simulation for Synthetic Theaters of War

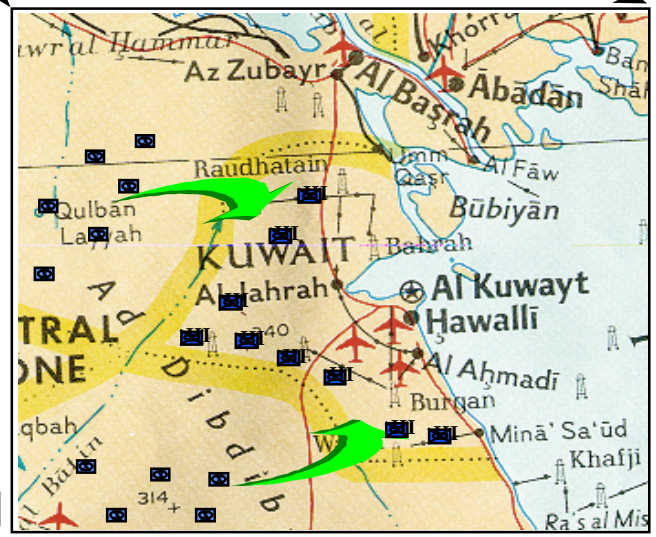


Achievements

- Entities 10,174 Nov. 1996
  30,000 May 1997
  50,000 June 1997
- Metacomputing for Combat Simulation
- Scalable Synthetic Forces Architecture
- Scalable Synthetic Environments Servers
- Gigabytes / minute data logging

Systems Environments' Technologies

- **Scalable I/O**
- **MPI-CH**



HPC Sites

ORNL	1024 Node Paragon
Cornell	400 Node SP2
Caltech/JPL	256 Node HP SPP2000
Ames	143 Node SP2

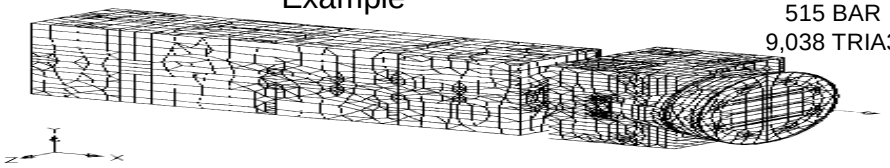
Networks

NREN, CERFNet, ESNet, vBNS, BAGNet

Parallelization of a Commercial Engineering Software

- A prototype of MSC/NASTRAN parallelized for linear statics, normal modes and frequency response solution sequences based on parallelizing numerical solution modules only.
- Started work on parallelization based on physical domain decomposition and parallelizing data recovery.
- Program is on schedule.
- Number of benchmark analyses performed for statics and normal modes and results verified by end-users.
- Performance targets met and even exceeded for static analysis on IBM SP2.
- The parallel efficiency for frequency response module is highly scalable. Preliminary numbers show about 85% parallel efficiency for 8 processors.

Example

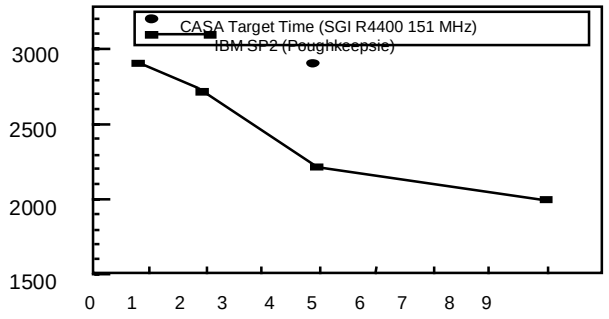


12,291 GRIDS
73,746 DOF
213 BEAM
10,309 ELAS1
6,009 QUAD4
515 BAR
9,038 TRIA3

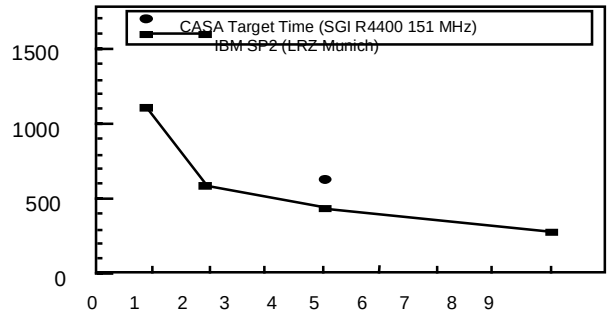
CASA Satellite Model - From Europort

Analysis	Target Time	Time With PMN	# Procs.	Machine
Static	643 sec	357 sec	8	(LRZ)
Modal	2882 sec	2026 sec	8	(Pou)

Lanczos Results



Statics Results



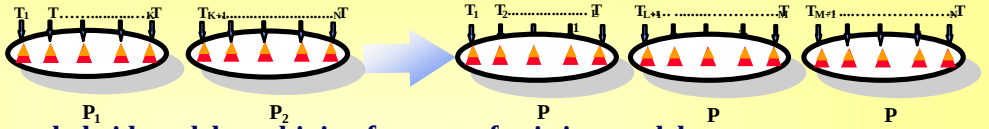
Application Programming System Challenges and Approaches

The present models poorly serve applications that:

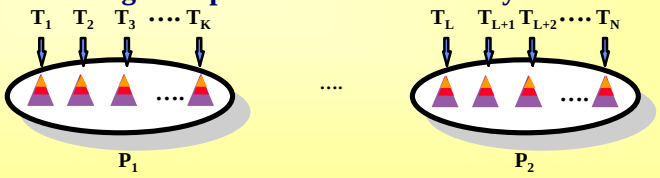
- **Require dynamic task scheduling and resource allocation**
  - **Tree-based algorithms (database searches) in MSTAR**
  - **DD applications, e.g. weather simulations using adaptive-mesh refinement**
- **Need task independence in the presence of multiple levels of memory hierarchy**
- **Must run on distributed heterogeneous platforms**

Research will overcome the limitations with a new model and compiler technology:

- **Extend SPMD, compiler directives, negotiation with QUORUM**
  - **Dynamically chunk the queue (ROI example)**



- **Develop a hybrid model combining features of existing models**
- **Create a new model allowing multiple levels of concurrency and data distribution**



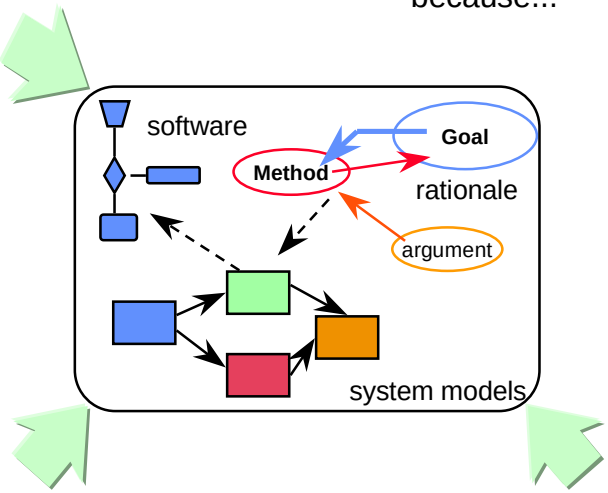
Evolutionary Design of Complex Software

Develop the technologies needed for continuous evolution of families of long-lived military software systems with costs proportional to the size of the change

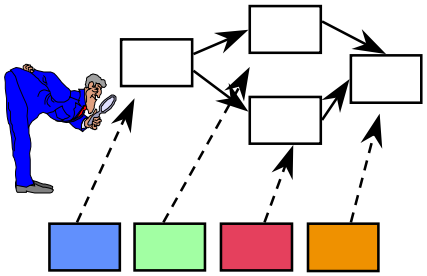
Design Management
Ensure that developers have access to all needed information and tools

Do it this way because...

That won't work because...



Architecture and Generation
Raise the level at which we think about systems



Test and Recertification
Demonstrate tools to recertify changed system with cost proportional to size of change

Environment

- Long system lifetimes
- Changing missions
- Stovepipe development
- Loss of design rationale
- "Maintenance" treated as an afterthought
- Languages & tools sacrifice flexibility for efficiency
- Commercial sector focus on high-volume, modest reliability & complexity, not DoD needs.

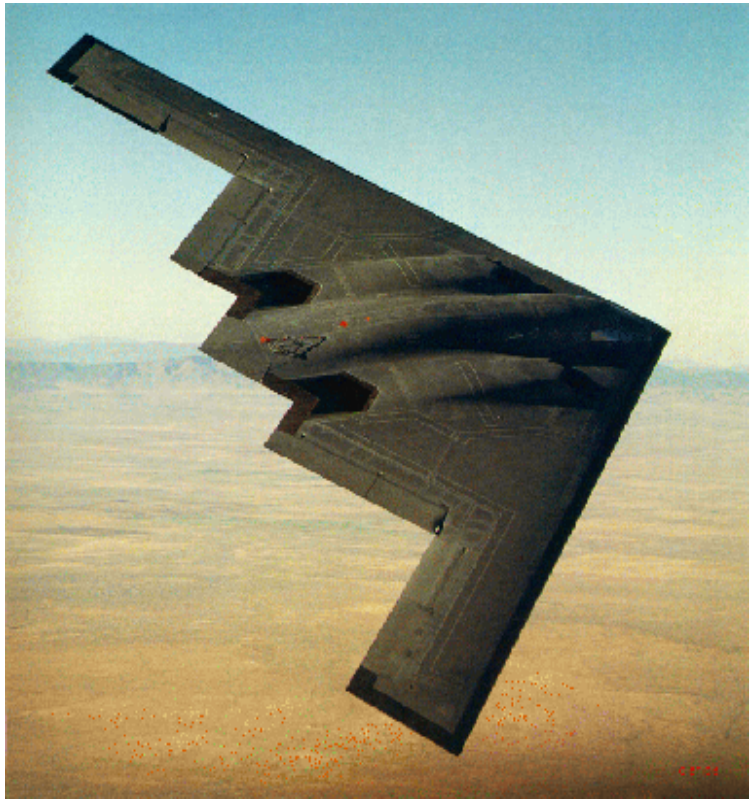
Goals

- Reduce the cost of changes and avoid errors by:
  - Providing more effective capture and management of information, tools, and products.
  - Allowing systems or system changes to be analyzed in more detail at all stages of design and efficient code to be generated (and regenerated) automatically.
  - Producing the tools needed to guarantee user-specified properties at both design and code levels, to greatly reduce the effort involved in regression testing, and to ensure safe backups for upgrades to high reliability critical systems.
- Demonstrate, measure and quantify the impact of integrated EDCS technologies in production software development settings.

B2 Bomber

B2 Bomber: Northrop/Grumman

- **Create a local-area network-based software engineering environment which simulates the B-2 software**
  1) Demonstrate rapid prototyping
  2) Demonstrate software understanding capabilities to support implementation of new functionalities
  3) Demonstrate incremental test and certification activities
- **Participants: Computing Services Support Solutions, Harlequin, Georgia Tech, UC Irvine, USC/CSE, USC/ISI**
- **Both intra-cluster and inter-cluster demonstration**



E-8C Joint Surveillance Target Attack Radar System (JSTARS)

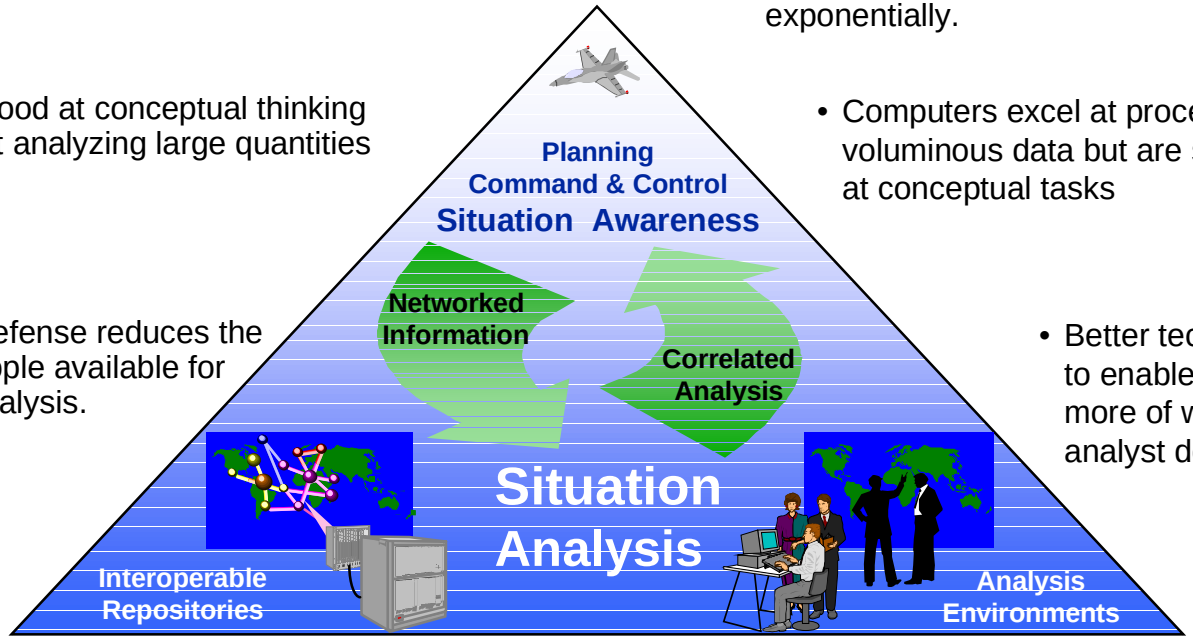
E-8C Joint Surveillance Target Attack Radar System (JSTARS)

- Apply to JSTARS system engineering activities within Northrop Grumman
  - Requirements negotiation and management
    - Change impact analysis
    - Legacy data integration
    - Intelligent documentation

  - CORBA-based infrastructure for system representation and analysis
  - System engineering tool development and COTS tool integration for Northrop Grumman

Information Management

- Information superiority depends on timely and accurate information discovery, filtering, and fusion.
- Humans are good at conceptual thinking but are poor at analyzing large quantities of information.
- Downsizing Defense reduces the number of people available for information analysis.
- But the number of document sources and the rate of document production continues to grow exponentially.
- Computers excel at processing voluminous data but are still relatively poor at conceptual tasks
- Better technology is required to enable computers to do more of what the human analyst does

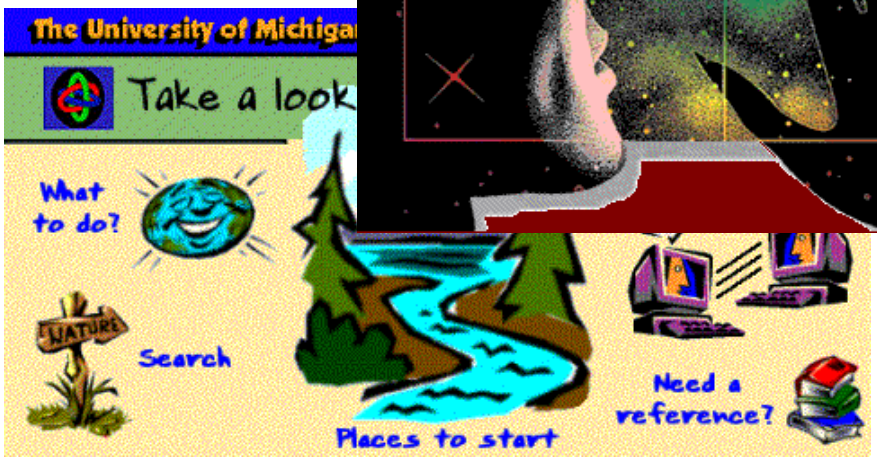


Make computers able to access, organize, analyze and disseminate the information contained in large, dynamic, multi-media document streams.

Applications in Life Long Learning

K-12 Education

- Learning Modules
- Photo Query by Content
- Searchable Video Skims



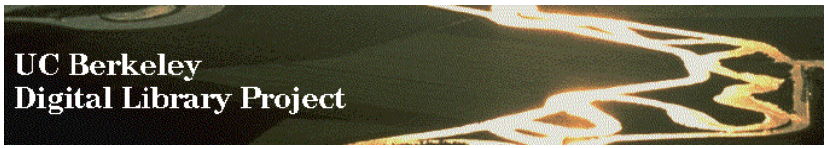
Continuing Education

- News on Demand
- Cross-discipline Navigation
- Dynamic Legacy Documents

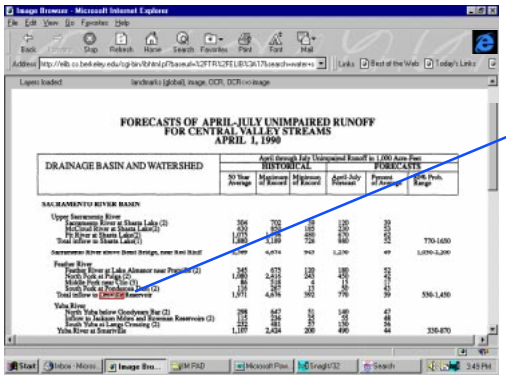
Applications in Crisis Management

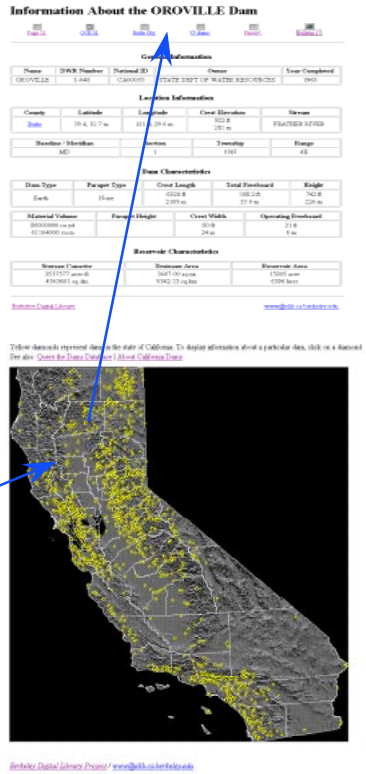
California Dept. of Water Resources (DWR) copes with floods of 1997

Oroville spillway - Normal



- Automated scanning & structuring of DWR reports and tabular data
- Near real-time situation updates
- Geographic, image, text, and numeric data





Oroville spillway - Flood

High-Performance Knowledge Bases:

What is a Knowledge Base?

Knowledge Base

Facts (DATA)
about the situation

Knowledge:
- Concept Definitions
- Laws, Constraints
- Relationships
- Domain Theories

Problem Solving Methods

Domain knowledge is explicitly stated in concept definitions, axioms & constraints.

Conventional Software

DATA

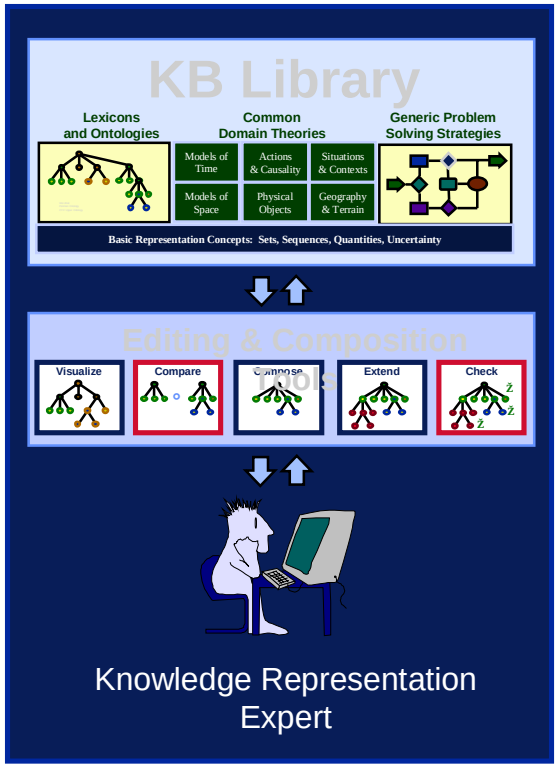
Procedural Code:
- Step-by-Step Instructions
- Optimized for computational efficiency

Domain knowledge is implicit in the programmer's step-by-step procedural instructions.

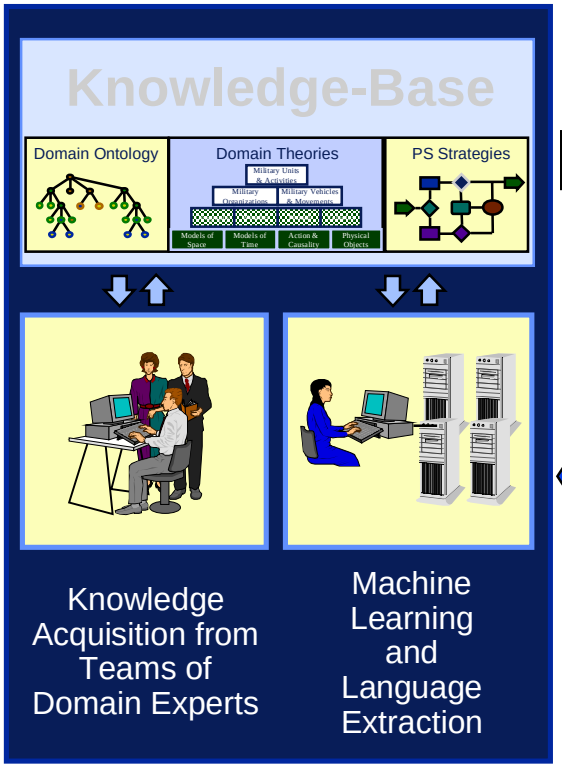
High-Performance Knowledge Bases:

Component Technology

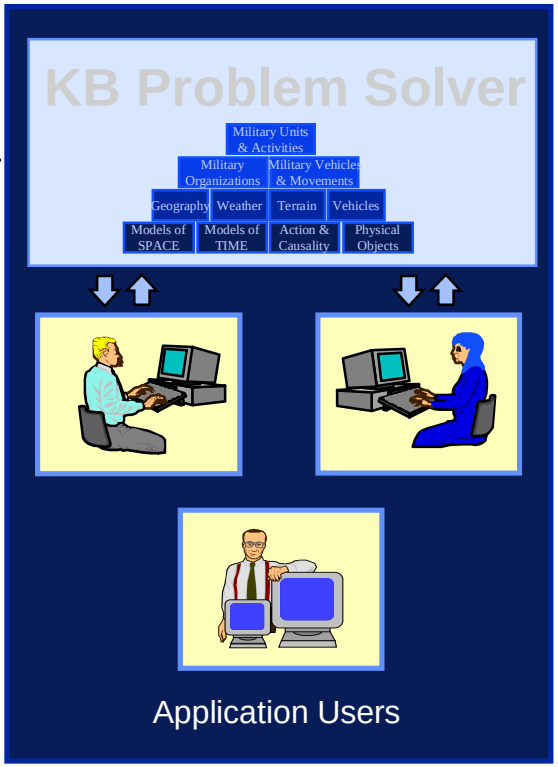
Foundation Building



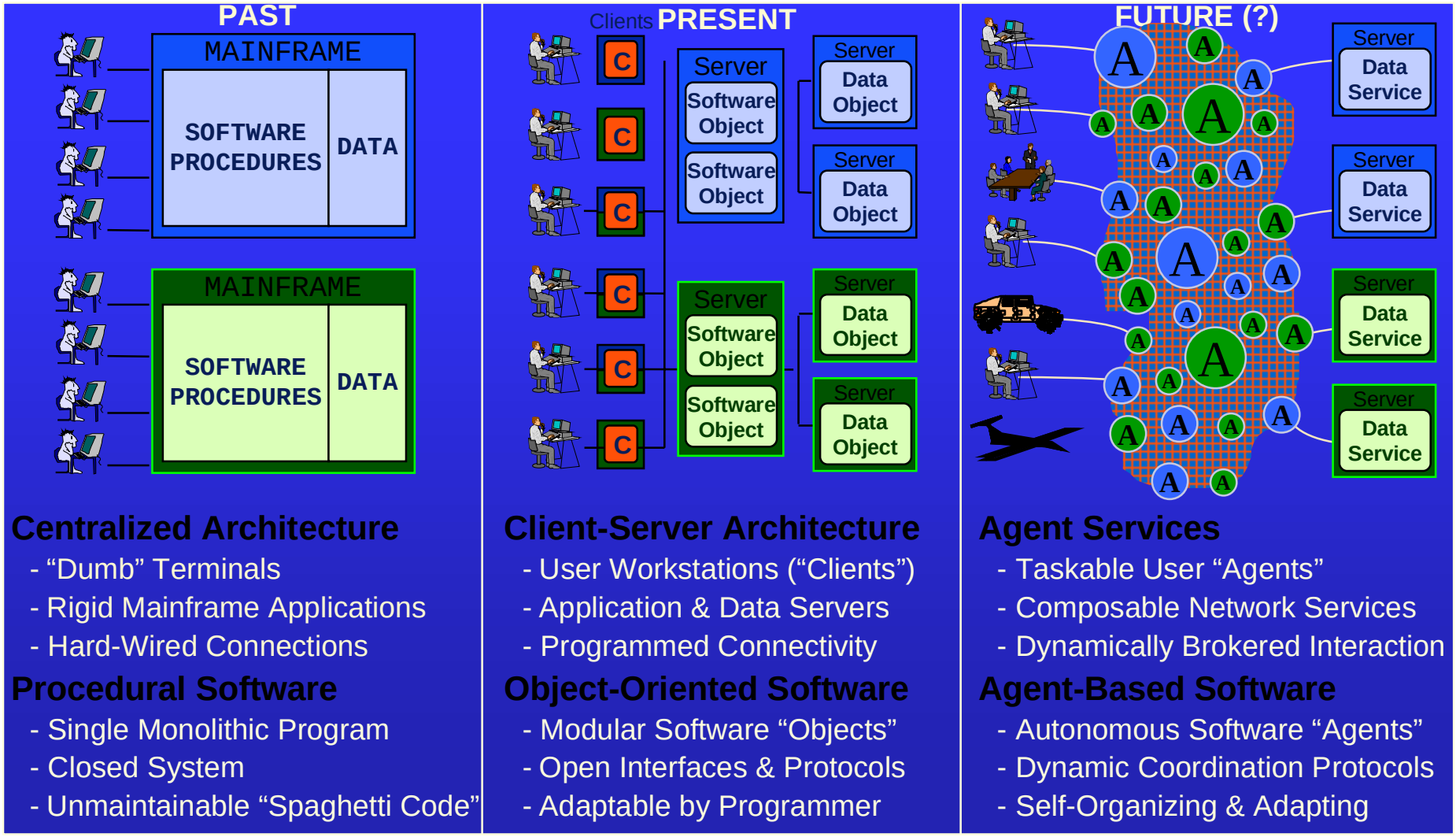
Knowledge Acquisition



Problem Solving & Use



Evolving Computing Environment



Centralized Architecture
- "Dumb" Terminals
- Rigid Mainframe Applications
- Hard-Wired Connections

Procedural Software
- Single Monolithic Program
- Closed System
- Unmaintainable "Spaghetti Code"

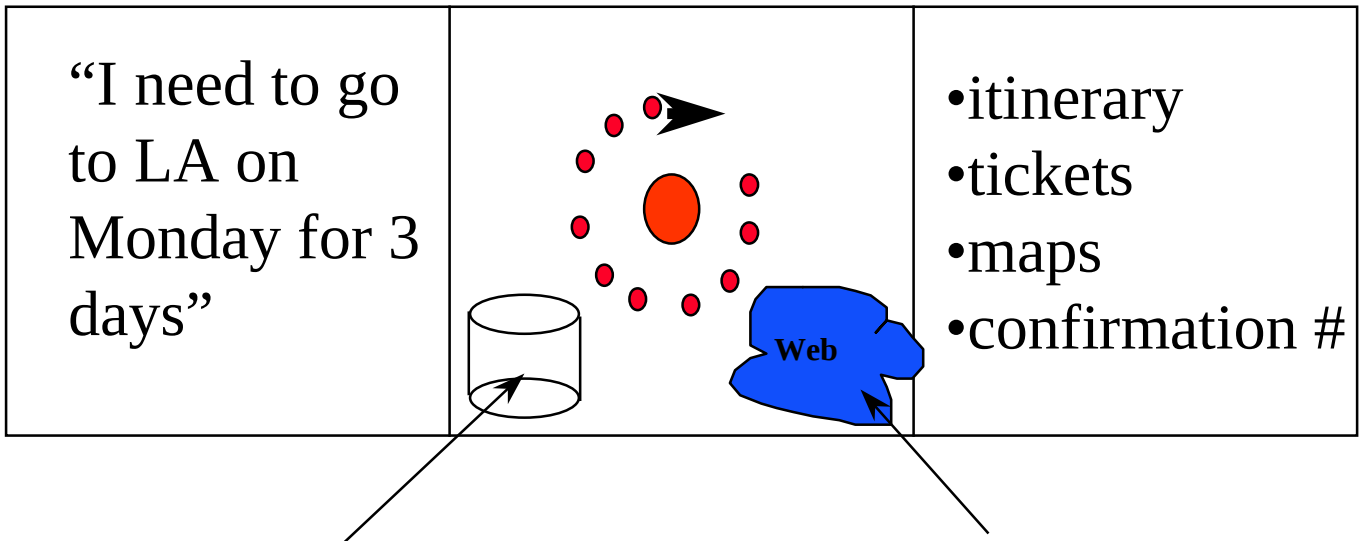
Client-Server Architecture
- User Workstations ("Clients")
- Application & Data Servers
- Programmed Connectivity

Object-Oriented Software
- Modular Software "Objects"
- Open Interfaces & Protocols
- Adaptable by Programmer

Agent Services
- Taskable User "Agents"
- Composable Network Services
- Dynamically Brokered Interaction

Agent-Based Software
- Autonomous Software "Agents"
- Dynamic Coordination Protocols
- Self-Organizing & Adapting

Simple example: travel “agent”



Credit card #
Preferences learned or supplied

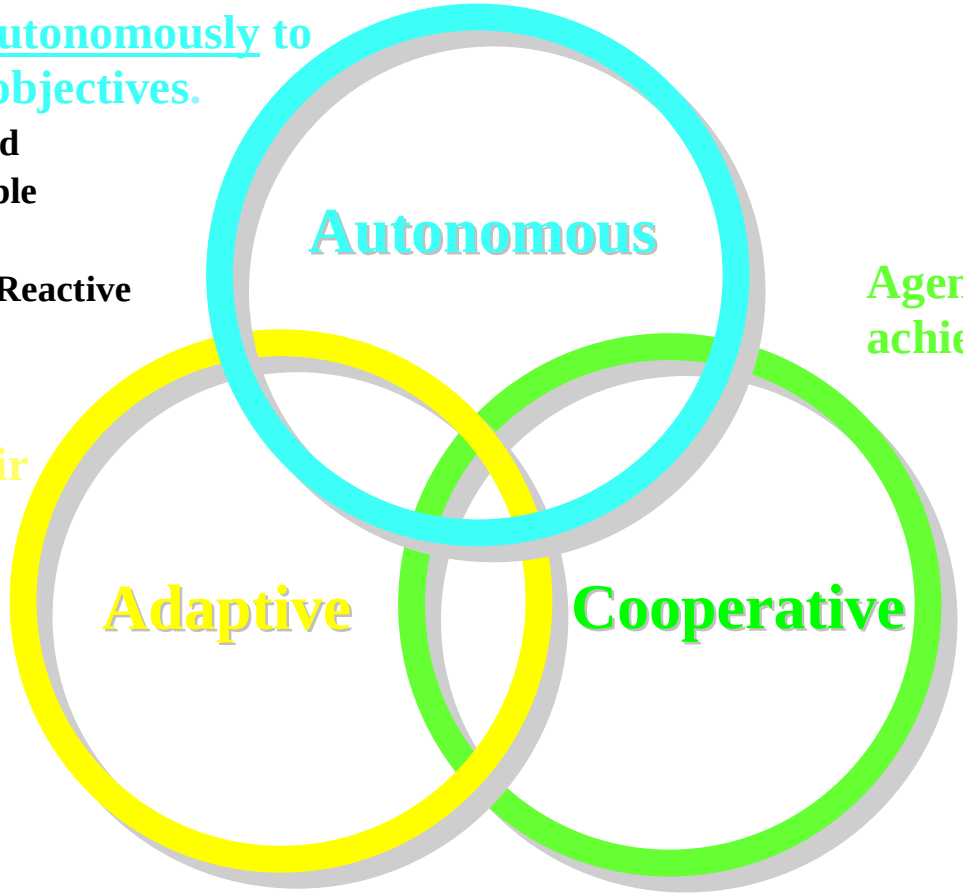
Web sites of prices for airlines, hotels, rental cars, and Federal Express

Biggest Problem: Will the agent buy undesired items?

Characteristics of Software Agents

Agents act autonomously to accomplish objectives.

- Goal-Directed
- Knowledgeable
- Persistent
- Proactive & Reactive



Agents cooperate to achieve common goals.

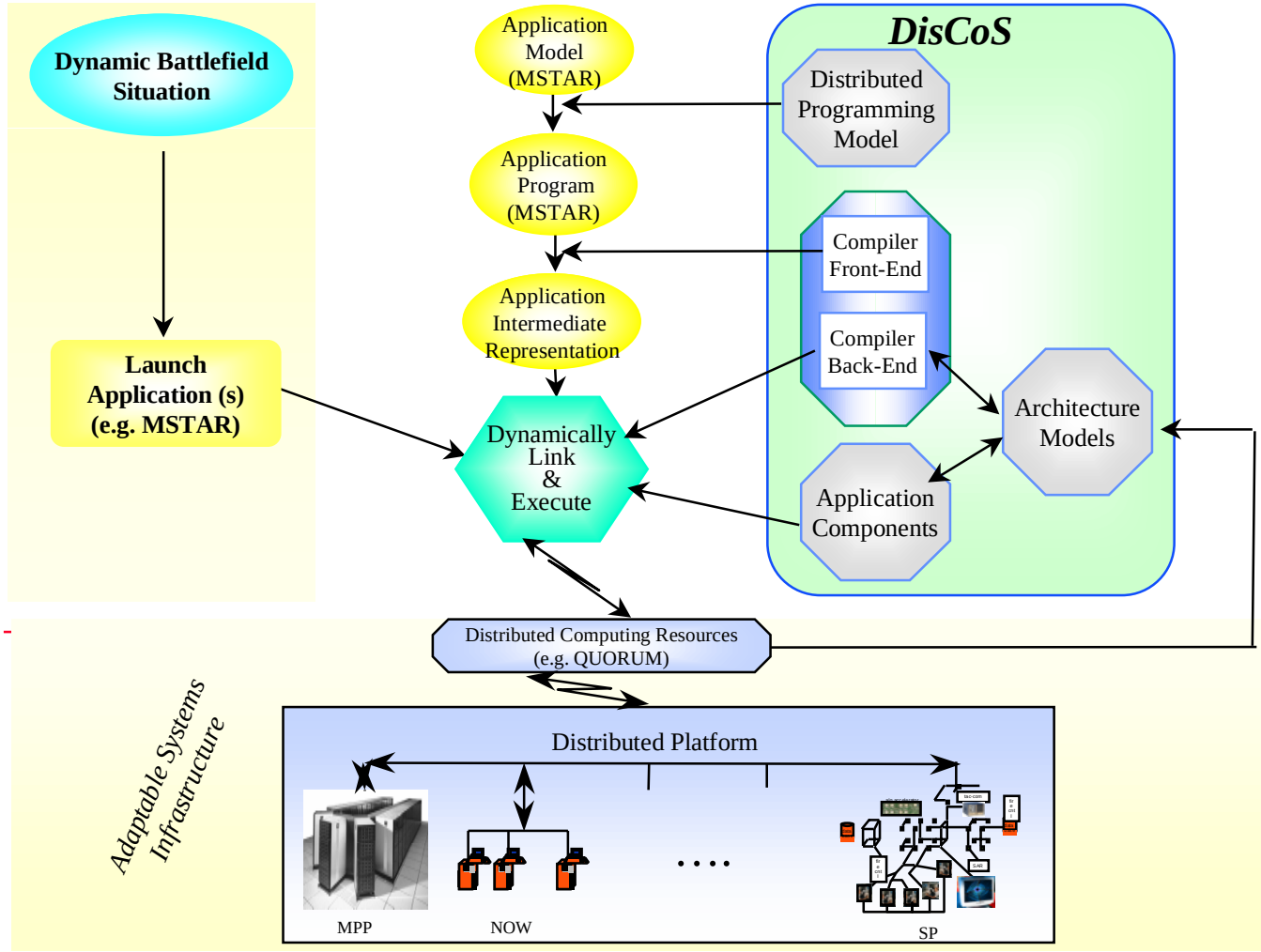
- Communication Protocols
- Knowledge-Sharing
- Coordination Strategies
- Negotiation Protocols

Agents adapt to their environment.

- Dynamic Interaction
- Alternate Methods
- Machine Learning

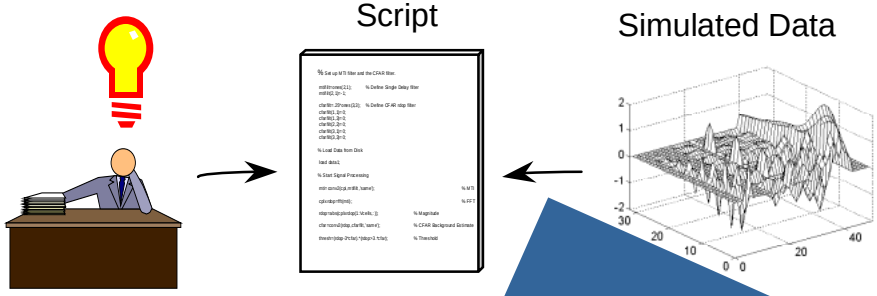
Note: Agents can be either static or mobile, depending on bandwidth requirements, data vs. program size, communication latency, and network stability

Exploiting Performance Monitoring



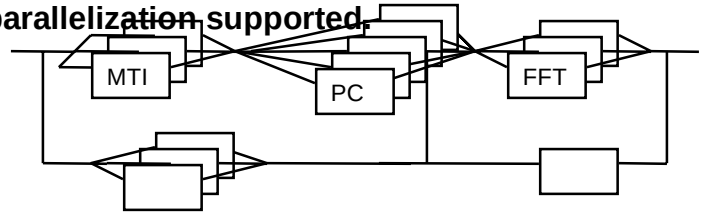
Realtime HPC MatLab User Perspective

1) Create script file, test and simulate algorithm on workstation



2) Add Real-time toolbox I/O functions to MatLab script. m-file emulation allows workstation simulation to continue.

3) Partition script into functional groups of MatLab script, assign nodes to groups using Target Balancing Tool. Pipeline, data parallel, task parallel, round robin parallelization supported.



4) Run "genmake". This creates a makefile for the machine architecture desired. Single & double precision math supported.

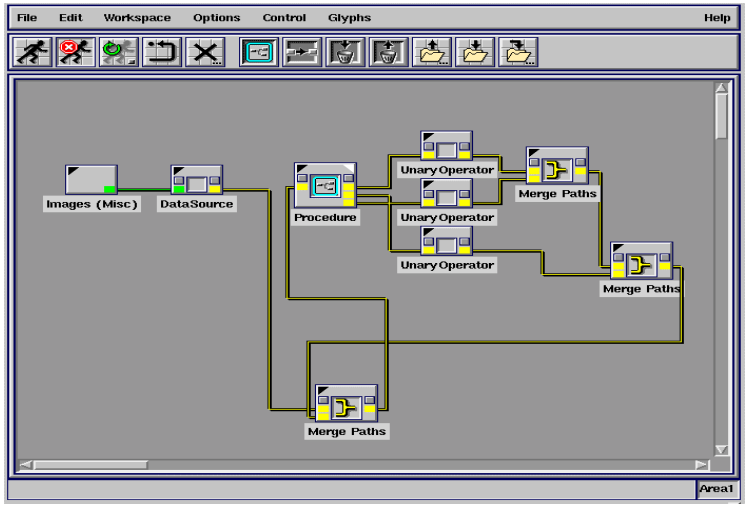
5) Run "make". The MatLab to C translation , C compile, and link is performed at this point.

Embedded Khoros

Khoros targeted for Embedded Realtime Systems

Target Systems

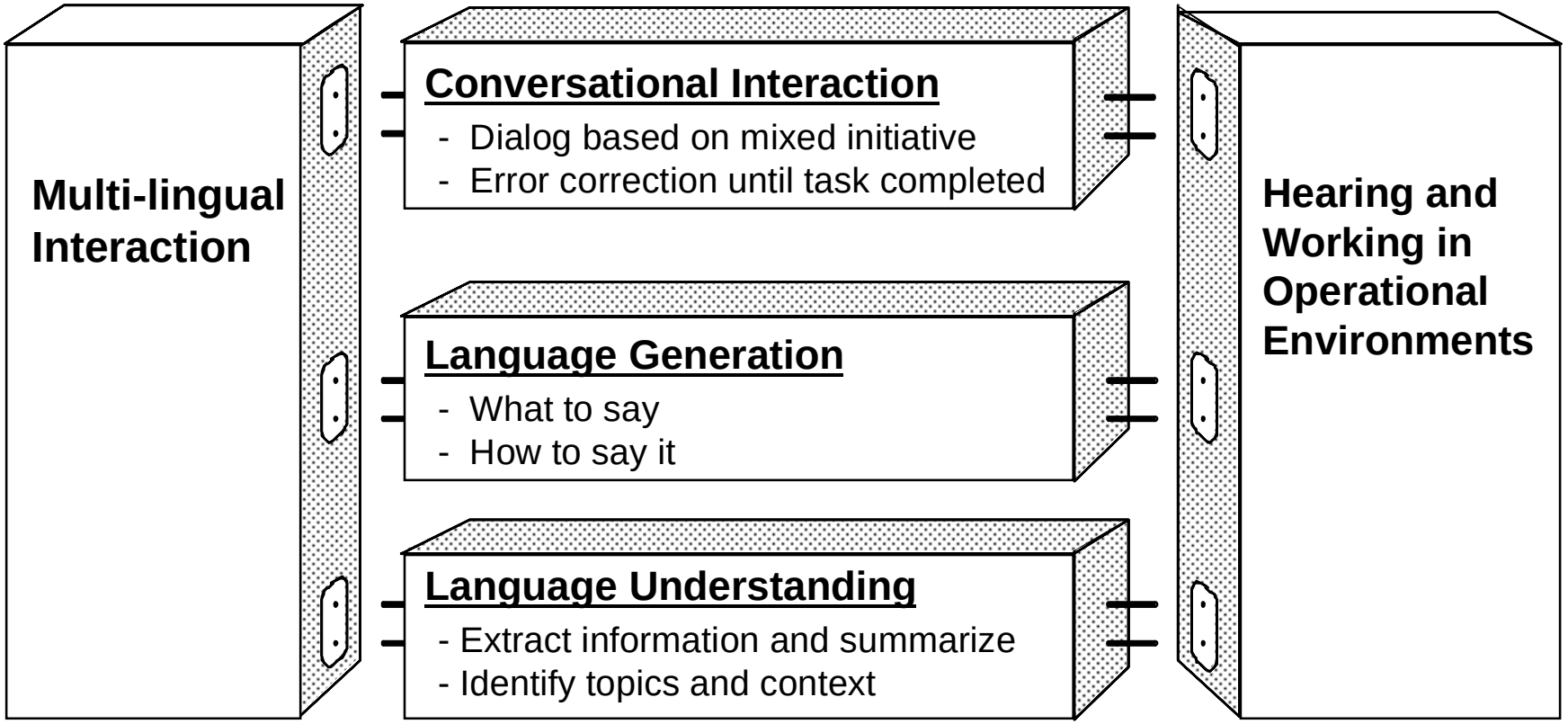
- Mercury Raceway
- Sanders HPSC



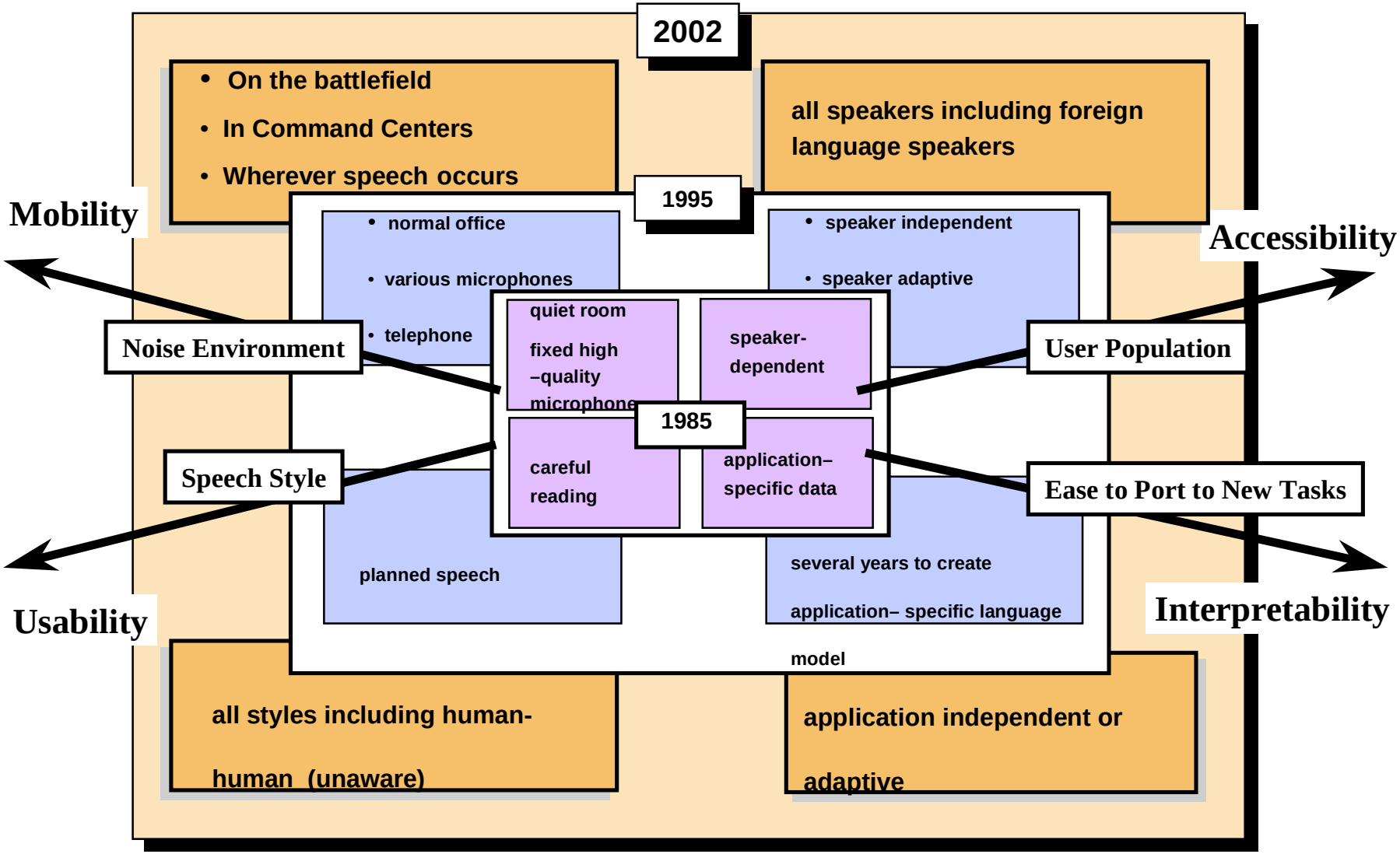
Parallel Machines

- IBM SP2
- Cray T3E
- Cambridge

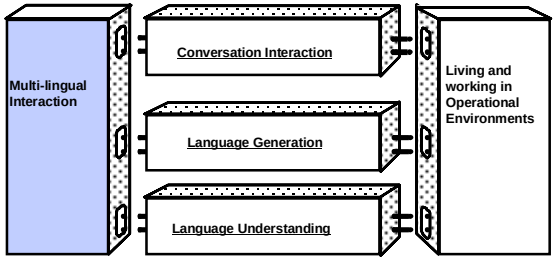
Technology Components

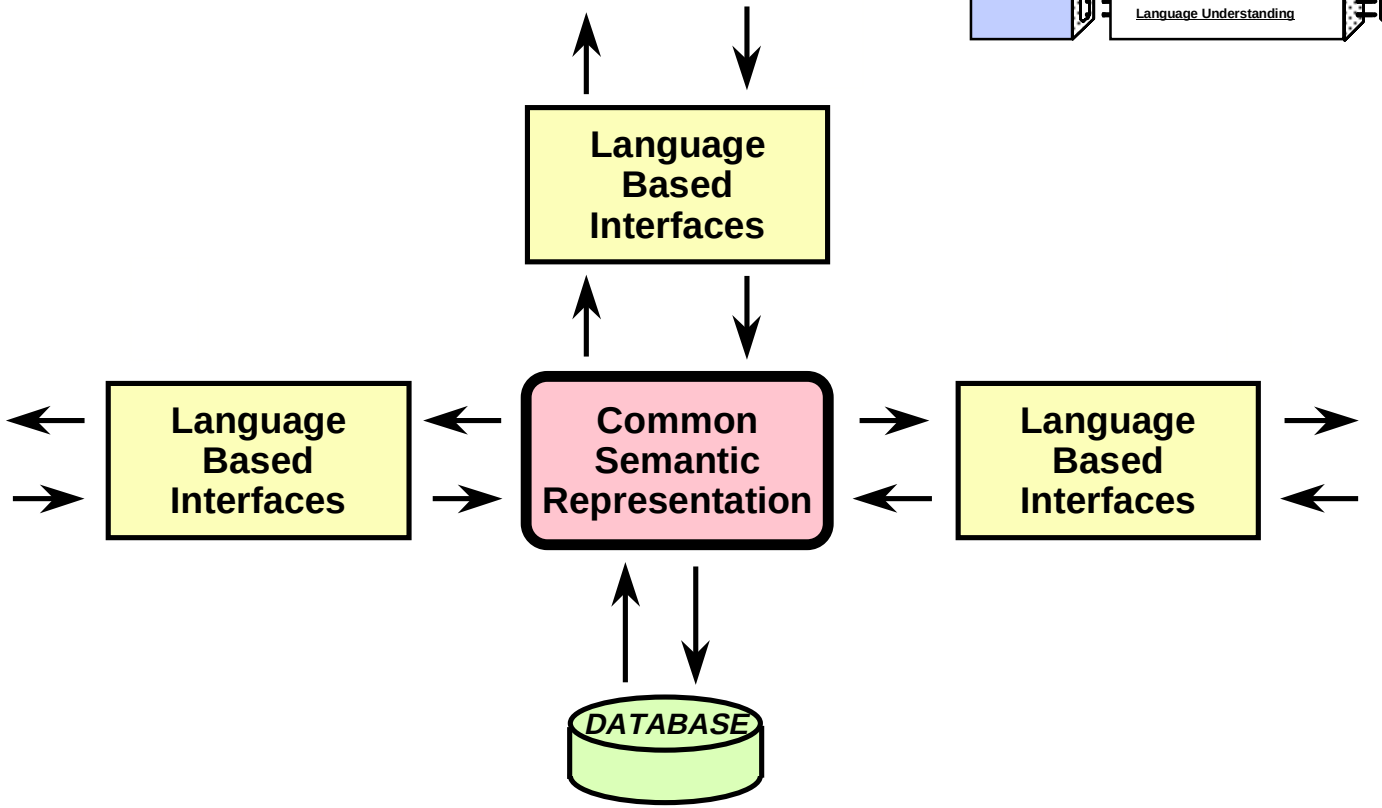


System Challenges for 2002



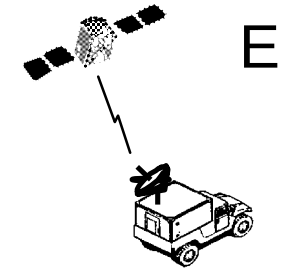
Multilingual Conversational Systems Human-Computer and Human-Human Interactions



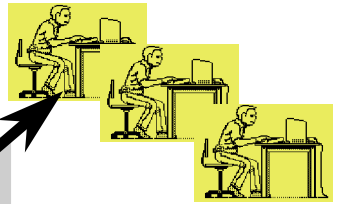


Human Language Systems

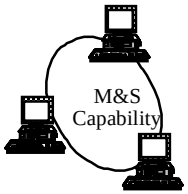
Enable computers to hear and read, as well as understand what they hear and read



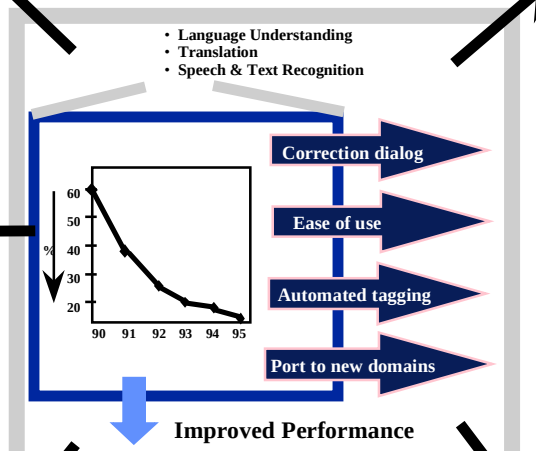
Mobile Computing



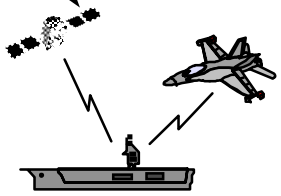
News-on-Demand



Distributed Simulation Set-Up and Control



Eyes and Ears of Log Anchor Desk

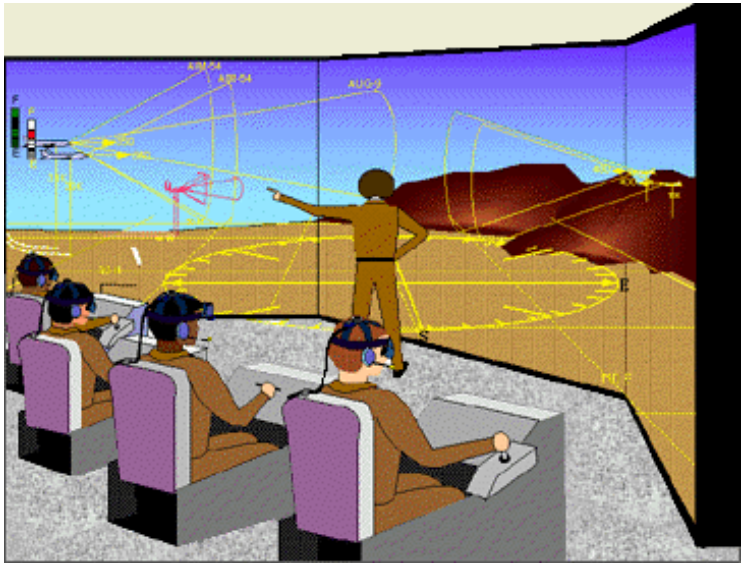


Joint Task Force Operations

Environment

- Information Overload ... too much info for analysts or users to process
- Information not processed fast enough
- Information in multi-languages
- Limited ability to identify language, topic and speaker
- Unaffordable software that does not adapt to user needs

CommandTalk



"Create an M1 platoon at 945 970"

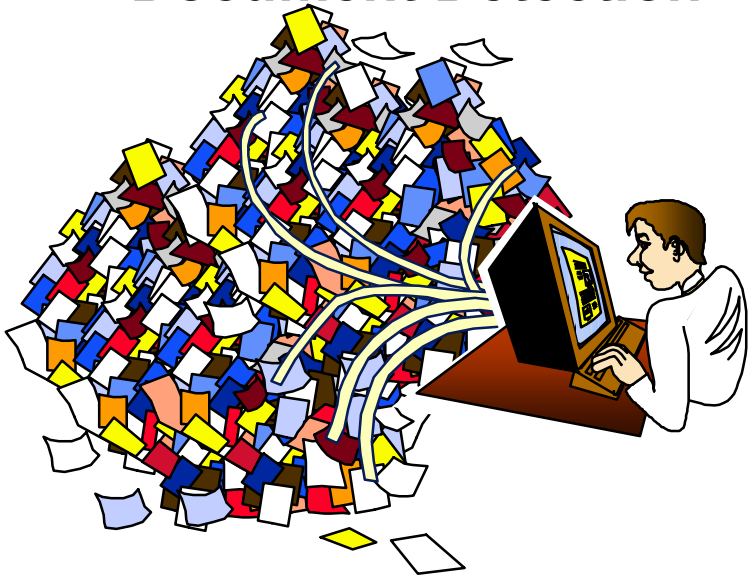
"First platoon, on my command, advance in a column to Checkpoint 1"

- Spoken language interface to synthetic forces in entity-based simulations
- Selected for use to control all forces (Army, Navy, Air, Marine Corps) in STOW-97
- Spoken language understanding technology:
  - Faster, more natural interface
  - More realistic training
  - High accuracy speech

TIPSTER

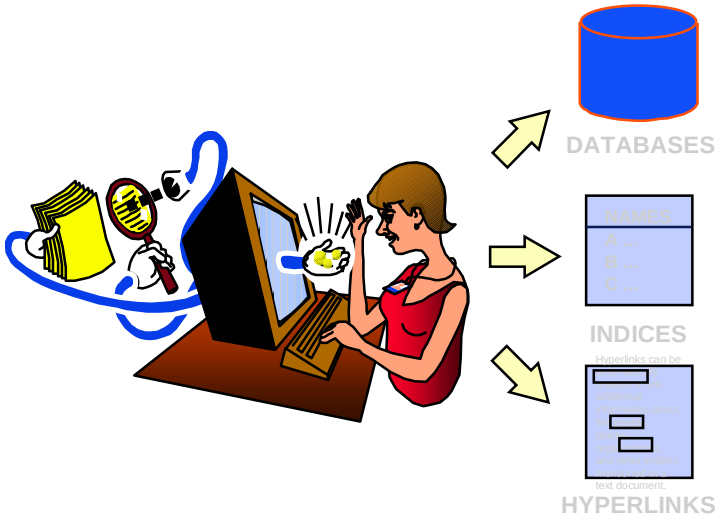
Text Document Understanding Initiative

Document Detection



- Document Routing
- Ad Hoc Retrieval
- Improved Accuracy
- Retrieval of Error-Filled Data
- Multilingual Retrieval

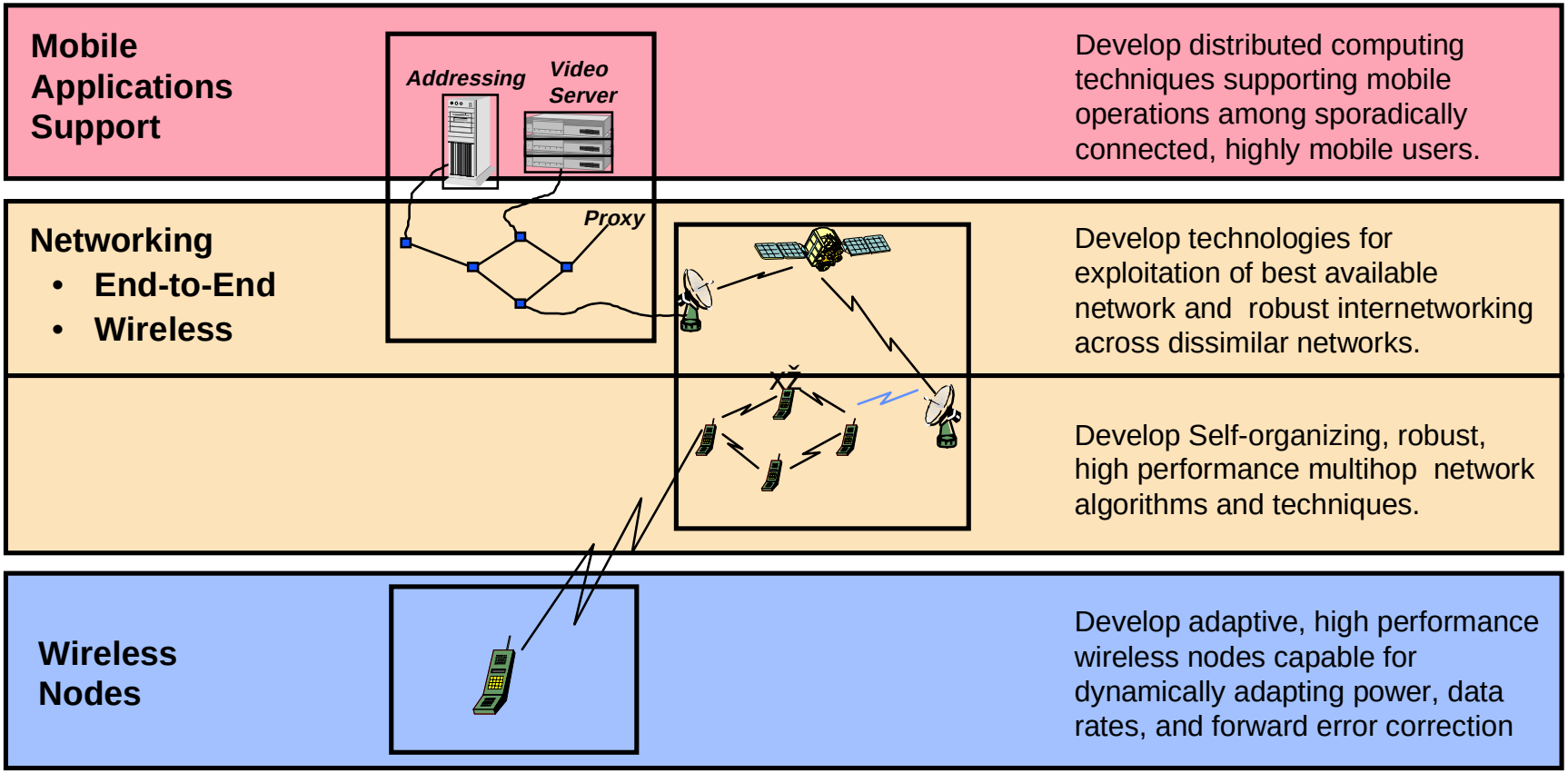
Information Extraction



- Extracting Named-Entities
- Extracting Relationships
- Extracting Events
- Extracting Scenarios
- Summarizing Documents

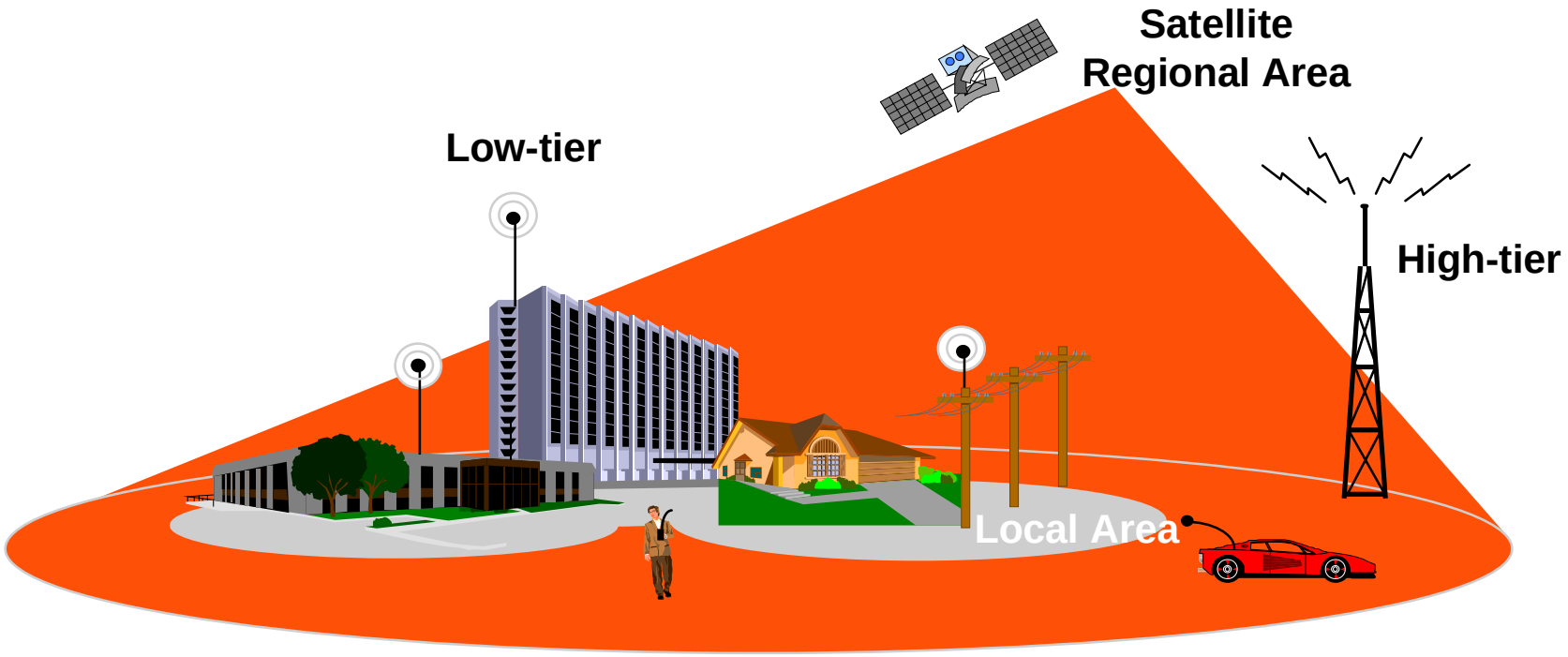
Untethered Nodes

- Enable technology for robust end-to-end information systems in a global, mobile environment
- Develop technology for integration of underlying commercial components into flexible, robust, multihop, high bandwidth system
- Demonstrate in realistic applications environment



Commercial Scenario

To enable users to take their "InfoEnvironment" with them …



Global Mobile Information Systems

Mobile Application Support

Support for sporadic and disconnected services	CMU
Mobile computing infrastructure software - includes peer to peer data base replication & reconciliation	Platinum (formerly Locus)
Support varying bandwidth of network connections to include video transcoding	UCB (BARWAN)
Location Transparent Computing relocation and data services	UCLA (TRAVLER)
Dynamic negotiation between application and network for quality of service	CMU (PYXIS)
Information Stations and Geo Services	Rutgers (NIMBLE)